On the joint dynamics of network diameter and spectral gap under node removal

Klaus Wehmuth, Antonio Tadeu A. Gomes, Artur Ziviani, Ana Paula Couto da Silva

On the joint dynamics of network diameter and spectral gap under node removal

Klaus Wehmuth, Antônio Tadeu A. Gomes, Artur Ziviani
National Laboratory for Scientific Computing (LNCC)
Petrópolis, RJ – Brazil
Email: {klaus,atagomes,ziviani}@lncc.br

Ana Paula Couto da Silva
Federal University of Juiz de Fora (UFJF)
Juiz de Fora, MG – Brazil
Email: anapaula.silva@ufjf.edu.br

***Abstract*—The study of complex networks, their properties, and underlying models has been the focus of a lot of attention in recent years. In particular, the network diameter and spectral analysis have been used to characterize the size and the robustness of such networks, respectively. In this work, we analyze the joint dynamics involving the behavior of the network diameter and the spectral gap under different node removal processes, showing their correlated behavior during the process.**

I. Introduction

Since the first findings that many complex networks did not follow a random structure as previously believed [1], [2], a lot of research has been performed in analyzing the structure and properties of different complex networks, leading to new ways of describing and measuring such networks [3]–[5]. A particular metric of interest in complex networks is the network diameter, defined as the longest shortest path between any pair of nodes composing the network. Such a metric is particularly important as it has been shown to be surprisingly small in comparison to the size of the network for many complex networks (see for instance [6]), a consequence of the small-world [1] and scale-free [2] properties of such networks. In parallel, spectral analysis [7] has been used to analyze the connectivity level of networks as a way to evaluate their robustness to an eventual partition [8]–[11].

In this work, we analyze the joint dynamics involving the behavior of the network diameter and the spectral gap under different node removal processes, showing their correlated behavior during the process. Although it is known that the spectral gap reflects the connectivity level of the studied graph, usually only the extreme cases (almost fully connected or likely to partition) are shown as examples. Thus, we focus on investigating how the spectral gap evolves as the network connectivity degrades under different node removal processes—ranging from ones representing random faults to ones corresponding to strategic attacks on the network structure—as well as how this impacts the network diameter.

This paper is organized as follows. In Section II, we briefly review the basic spectral analysis properties needed for the presented study. We present our analysis on the joint dynamics of network diameter and spectral gap under node removal in Section III. In Section IV we provide some concluding remarks and perspectives for future work.

II. Background

A graph is defined as an ordered pair (V, E) where V is the set of vertices and E is the set of edges. An edge is a set $\{v1, v2\}$ where $v1, v2 \in V$. This definition is sufficient for representing undirected graphs with at most one connection between any given pair of different vertices, and no connection from a vertex to itself.

The Laplacian Matrix $\mathbf{L}$ is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{A}$ is the adjacency matrix of the considered graph and $\mathbf{D}$ is the diagonal degree matrix of such a graph, *i.e.* the value of element (n, n) in $\mathbf{D}$ corresponds to the degree of the n^{th} vertex in (V, E). Clearly the vertex sequence on the representation of $\mathbf{D}$ and $\mathbf{A}$ needs to be the same. From the definition of the Laplacian Matrix $\mathbf{L}$, it can be seen that the sum of all columns of $\mathbf{L}$ is the zero vector $(\vec{0})$. Therefore, it follows that $\mathbf{L}$ is singular and that the vector $(\vec{1})$ is in the *Nullspace* of $\mathbf{L}$. As a consequence, 0 is an eigenvalue of $\mathbf{L}$ and $(\vec{1})$ in an eigenvector associated to the eigenvalue 0.

For a connected graph, the Normalized Laplacian Matrix $\mathcal{L}$ is defined as $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. Alternatively, $\mathcal{L}$ can also be defined as

$$\mathcal{L}_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ \frac{-1}{\sqrt{d_i d_j}} & \text{if } i \text{ is adjacent to } j, \\ 0 & \text{otherwise,} \end{cases}$$

where d_k is the degree of the vertex k.

The Normalized Laplacian Matrix $\mathcal{L}$ has some properties that are of particular interest in this work. First, independent of the network size, all eigenvalues of $\mathcal{L}$ are between 0 and 2 due to the normalization procedure, *i.e.*, $0 = \lambda_1(\mathcal{L}) \leq \lambda_2(\mathcal{L}) \leq \cdots \leq \lambda_n(\mathcal{L}) \leq 2$. This is an important feature since our goal is to analyze the changes on the network structure caused by node removal, regardless of the change in the network size itself. Second, $\lambda_2(\mathcal{L})$, the smallest non-zero eigenvalue, is less than 1 if the graph is not complete, reflecting the graph connectivity level approaching 0 as the graph tends to be less connected. This particular eigenvalue, $\lambda_2(\mathcal{L})$, also known as the *spectral gap*, is extensively used in this work and it will be referred to simply as λ_2 hereafter.

III. Analysis

In this section we show our analytical results concerning the behavior of network diameter and the spectral gap in face of node deletion. Before that, however, we present the set of tools used to create and analyze the networks of interest.

A. Tools

We use the following tools in this study:

- Networkx [12] is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. This tool is used to generate random graphs, as well as graphs specifically defined from a list of vertices and edges. After the graphs are created, this tool is also used to manipulate these graphs by adding or removing vertices or edges.
- NumPy [13] is a Python library for working with multi-dimensional arrays.
- SciPy [14] is is open-source software for mathematics, science, and engineering. The algorithm of choice for calculating eigenvalues is provided by SciPy on its sparse matrices linear algebra module. The algorithm implemented in this module is LOBPCG [15].

B. Analytical results

We consider three ways of removing nodes from a graph and applied them to two different kinds of graphs. The first node removal method is a strategic removal, where the node with highest degree on the graph is removed first. If more than one node is found with the same highest degree, one of the nodes sharing the highest degree is randomly chosen and removed. This deterministic node removal strategy represents strategic attacks on the network structure targeting the most connected nodes. The second node removal method is random with a degree bias. In this case, a node is chosen to be removed with a probability proportional to its current degree. The third removal method is uniformly random. In this case, a node is randomly chosen and any node is equally likely to be chosen. The uniformly random strategy represents random faults across the network structure. These three node removal methods are applied to both Barabási-Albert (BA) [2] (*i.e.*, scale-free) and Erdös-Rényi (ER) [16] (*i.e.*, random) graphs.

The BA and ER graphs used in this work are generated using the Networkx tool. The BA graphs are generated giving a target number of nodes and a given number of edges to create for each new node. The BA graphs then were created with 1000 nodes and 2 link attachments per each new node, resulting in BA graphs with a volume of 3992 and mean degree of 3.992. The ER graphs were created with 1000 nodes using a 0.0045 probability for edge creation between each pair of nodes, thus resulting in graphs with 1000 nodes, a volume of approximately 4350 and mean degree of approximately 4.35.

Several BA and ER graphs with the described properties were created and submitted to the three different node removal procedures described above. For each node removed, the resulting graphs were analyzed for connectivity. In the case the graph resulting from the node removal was no longer connected, the largest component from the resulting connected components was kept and the others set aside. The following parameters were then calculated and logged: identity and degree of the removed node, volume, order, diameter, radius and normalized λ_2 for the remaining graph. This procedure was repeated until the diameter of the remaining graph has reached its maximum. Beyond this point, the graph was considered degraded and the node removal process interrupted.

In addition to the logged values, the graph and specific parameters of each experiment such as random seeds were also stored, so that each experiment could be repeated and if desired the resulting graph at any given step recovered. This was later used for detailed analysis of the effect of a single node removal by comparing the graphs before and after a particular node removal of interest, as well as any connected component generated by this specific removal. This was instrumental in giving a better understanding of causes of the dynamic behavior of the logged parameters.

In Figure 1, we have an example of the result obtained from a BA graph under a random node removal procedure. The solid curve using the scale on the left shows the value of λ_2, while the dashed curve with scale on the right shows the graph's diameter as nodes are randomly removed.

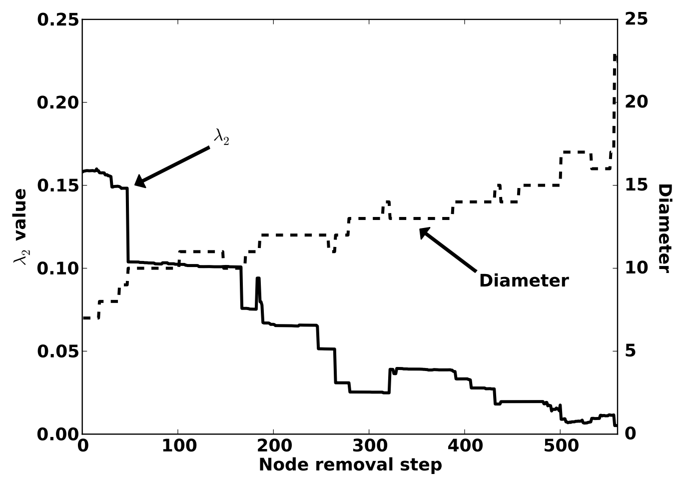


Fig. 1. Behavior of λ_2 and graph diameter under random node removal.

In this work, we focus on analyzing observed relationships in the behavior of the spectral gap λ_2 and the graph diameter. Different than the number of node removals needed to degrade the graph, which varies with the kind of graph and the adopted node removal method, as shown in [17], the dynamic behavior of λ_2 and the graph diameter shows similarities on both kinds of graphs and under different removal methods used in this work. Further analysis exposes these similarities as well as the differences encountered in each case.

From the observation of multiple instances of graphs similar to Figure 1, we have noticed that there is indeed some relation between variations in λ_2 and variations in graph diameter caused by node removal. There were cases where abrupt changes in λ_2 were associated with changes in diameter—an intuitive assumption—and cases where this does not happen—a rather counter-intuitive case. Analyzing in detail the effects of node removal associated with changes in λ_2 or network diameter, it could be seen that in general a significant drop in λ_2 is associated with a structure being formed where a small subgraph is connected to the main graph by a single node. When the removed node also used to create a shortcut in the graph structure, an increase in network's diameter can be observed simultaneously to a fall in λ_2. An abrupt increase in λ_2 is in general associated with the removal of a node that breaks away a subgraph structure similar to the one cited above. When this subgraph structure contains the path that determines the graph's diameter, the removal of the whole structure will cause the diameter to decrease.

In order to analyze the relationship between changes in the spectral gap λ_2 and the network diameter, we plot these changes in dispersion graphs to compare the variations on each parameter for each step in the considered node removal process. We submit seven BA and seven ER graphs, created with the parameters mentioned before, to the three kinds of node removal processes. The results are used to generate six dispersion graphs correlating the changes on λ_2 horizontally and the changes in graph diameter vertically for each combination of kind of graph and removal process. These dispersion graphs are shown in Figures 2(a) to 2(f). To make it possible to identify superposition of correlates in each dispersion graph, each correlate is drawn as a gray transparent circle. Thus, the darker a point is in a graph the more dispersion correlates there are at that point.

In Figure 1 we observe that most of the node removal steps cause rather small variations to λ_2 and no variation to the network diameter. This joint behavior of λ_2 and the network diameter is reflected on its corresponding dispersion graph (Figure 2(a)) as an accumulation of points on the horizontal axis close to zero. Also, as expected, it can be seen that λ_2 varies in a continuous way while the diameter varies in a discrete way.

Comparing these dispersion graphs, we observe that the relationship between λ_2 and network diameter follows particular patterns, showing similarities as well as differences for each of the six types of carried out experiments. For instance, comparing Figures 2(a), 2(c), and 2(e), which show the behavior of BA graphs under random, weighted random, and strategic node removal processes, respectively, we observe that, as stated in [17], this kind of graph is less sensitive to random removals than to strategically minded removal processes. The pattern in Figure 2(a) shows smaller changes on the network diameter during the node removal process, reflecting the less sensitive behavior of a BA graph to this kind of node removal process. It is also interesting to note that it is enough to have a strategically weighted random removal process—*i.e.*, the chance for a node to be chosen for removal is proportional to its degree—to achieve a result similar to a deterministic strategic node removal where the highest degree node is always chosen to be removed first. Comparing Figures 2(b), 2(d), and 2(f), we observe that, as expected, an ER graph is less sensitive to the application of different node removal algorithms because the patterns shown in these figures are similar to each other.

Focusing our attention on the lower-right part of Figures 2(a) to 2(f), we observe a distinct behavior for each kind of node removal process. In Figures 2(a) and 2(b), when a random removal process is considered, there is a higher concentration of dispersion correlates on the lower-right quadrant of the graph than for other node removal processes shown in the other subfigures. A medium concentration of dispersion correlates is observed when using random-weighted removal in Figures 2(c) and 2(d), while the lowest concentration is observed when using a strategic node removal, as shown in Figures 2(e) and 2(f). This behavior is consistent for both BA and ER graphs, indicating a trend when considering the different node removal processes. This indicates that the more the node removal process is strategically minded, the less likely it is to have a network diameter decrease. This shows that the destruction of the appendix-like subgraph structure usually related to these events is most likely associated with the removal of a low degree node, since, in general, this is the kind of node that connects the appendix subgraph to the main graph component. It is also clear that increases in network diameter associated with increases on λ_2 are quite uncommon as well as decreases in network diameter associated with decreases of λ_2.

IV. Conclusion

In this work, we have shown some preliminary results on investigating the joint dynamics involving the behavior of the network diameter and the spectral gap under different node removal processes. Although it is known that the spectral gap reflects the connectivity level of the studied graph, usually only the extreme cases (almost fully connected or likely to partition) are shown as examples. With the present study, we shed some light on how the spectral gap behaves as the network connectivity degrades, including its joint dynamics with respect to network diameter. This is done under node removal strategies that represent a range from random node faults to strategic attacks on the network structure. Therefore, we provide insights to the structural changes imposed to the graphs by the considered node removal processes.

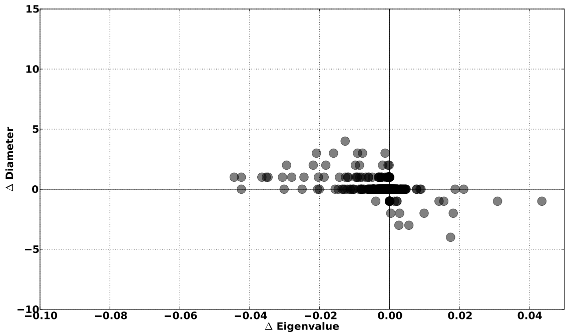
As future work, we intend to include refinements to the eigenvalue calculating algorithm, to perform further analysis of the behavior of the considered graphs under node removal—in particular, detailed analysis of counter-intuitive cases—as well as to consider other kinds of graphs, either synthetic or real-world ones.

References

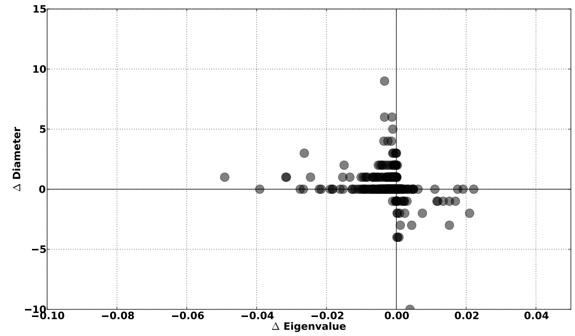
[1] D. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. June, pp. 440–442, 1998.

[2] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

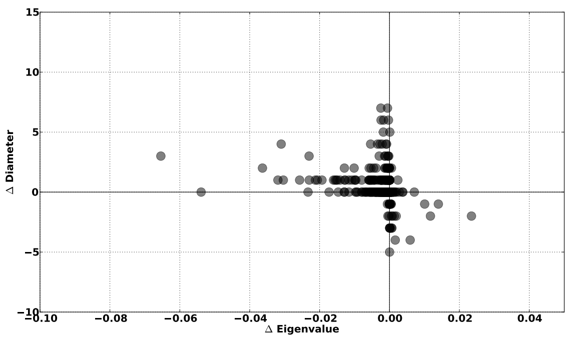
[3] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, pp. 47–97, 2002.



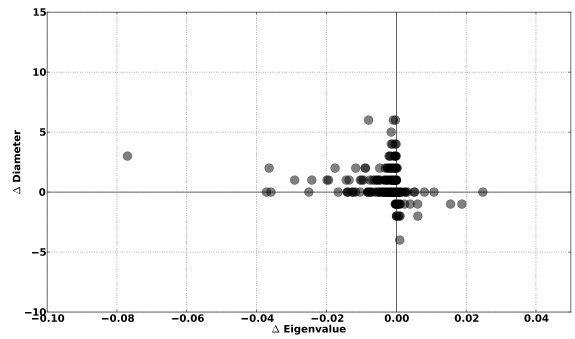
(a) BA graph under random node removal.



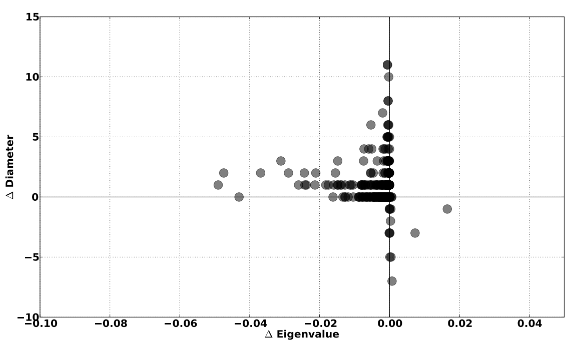
(b) ER graph under random node removal.



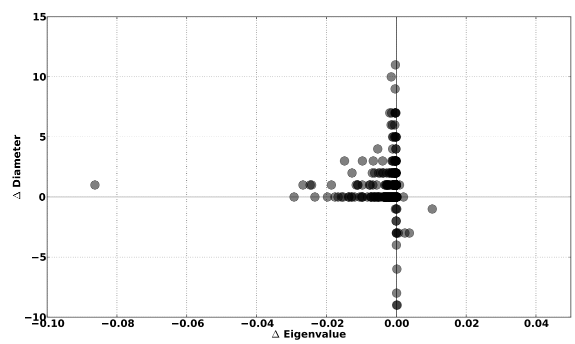
(c) BA graph under weighted random node removal.



(d) ER graph under weighted random node removal.



(e) BA graph under strategic node removal.



(f) ER graph under strategic node removal.

Fig. 2. BA and ER graphs under different approaches for node removal.

[4] M. E. J. Newman, "The Structure and Function of Complex Networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

[5] A.-L. Barabási, "Scale-Free Networks: A Decade and Beyond," *Science*, vol. 325, no. 5939, pp. 412–413, 2009.

[6] R. Albert, H. Jeong, and A.-L. Barabási, "Diameter of the World-Wide Web," *Nature*, vol. 401, p. 130, Jul. 1999.

[7] F. Chung, "Lectures on spectral graph theory," *CBMS Lectures, Fresno*, 1996.

[8] C. Gkantsidis, M. Mihail, and E. Zegura, "Spectral analysis of Internet topologies," in *IEEE INFOCOM*, vol. 00, no. C. IEEE, 2003, pp. 364–374.

[9] M. Mihail, C. Papadimitriou, and A. Saberi, "On certain connectivity properties of the internet topology," *Journal of Computer and System Sciences*, vol. 72, no. 2, pp. 239–251, Mar. 2006.

[10] A. Jamakovic and S. Uhlig, "On the relationship between the algebraic connectivity and graph's robustness to node and link failures," in *Proceedings of the 3rd EURO-NGI Conference on Next Generation Internet Network*. Trondheim, Norway: IEEE, May 2007, pp. 96–102.

[11] D. Fay, H. Haddadi, A. Thomason, A. W. Moore, R. Mortier, A. Jamakovic, S. Uhlig, and M. Rio, "Weighted Spectral Distribution for Internet Topology Analysis: Theory and Applications," *IEEE/ACM Transactions on Networking*, vol. 18, no. 1, pp. 164–176, Feb. 2010.

[12] "Networkx." [Online]. Available: http://networkx.lanl.gov/

[13] "Numpy." [Online]. Available: http://numpy.scipy.org/

[14] "Scipy." [Online]. Available: http://www.scipy.org/

[15] A. V. Knyazev, "Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method," *SIAM Journal on Scientific Computing*, vol. 23, no. 2, p. 517, 2001.

[16] P. Erdös and A. Rényi, "On random graphs I," *Publ. Math. Debrecen*, 1959.

[17] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–82, Jul. 2000.